Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics e-ISSN: 2656-8624
Homepage: https://ijeeemi.org/; Vol. 7, No. 3, pp. 539-548, August 2025

RESEARCH PAPER **OPEN ACCES**

Improving Diabetes Prediction Using Feedforward Neural Networks with Adam Optimization and the SMOTE Technique

Arizha Wijaya Kusuma[1], Muhammad Itqan Mazdadi[1], Dwi Kartini[1], Andi Farmadi[1], Fatma Indriani[1], and Chandrasekaran, P.[2]

[1] Department of Computer Science, Lambung Mangkurat University, South Kalimantan, Indonesia
[2] K.S. Rangasamy College of Arts and Science (Autonomous), Tiruchengode, India

Abstract

Diabetes mellitus is a chronic metabolic disorder that demands early and accurate detection to prevent life-threatening complications. Traditional diagnostic procedures, such as blood glucose and oral glucose tolerance tests, are often invasive, time-consuming, and resource-intensive, making them less practical for widespread screening. This study explores the potential of artificial intelligence, specifically Feedforward Neural Networks (FNN), in predicting diabetes based on clinical data from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The main contribution of this research lies in the integration of the Adaptive Moment Estimation (Adam) optimization algorithm and the Synthetic Minority Oversampling Technique (SMOTE) to enhance the performance and generalization of FNNs on imbalanced medical datasets. The methodology involves preprocessing steps such as imputing zero values with feature means, normalizing input features using Min-Max scaling, and applying SMOTE to balance class distribution. Two model configurations were compared: a baseline FNN trained manually using full-batch gradient descent and an FNN optimized with Adam. Experimental results demonstrated that the baseline model achieved an accuracy of 70.13%, precision of 56.06%, recall of 68.52%, and an F1-score of 61.67%. In contrast, the Adam-optimized model achieved superior results with an average accuracy of 73.31%, precision of 60.97%, recall of 66.67%, and an F1-score of 63.64% across ten independent runs. These findings indicate that combining adaptive optimization with oversampling significantly enhances the robustness and reliability of neural networks in medical classification tasks. In conclusion, the proposed method provides a practical framework for AI-assisted early diabetes detection and highlights opportunities for future development using deeper network architectures and explainable AI models for clinical applications.

Paper History
Received June 10, 2025
Revised August 08, 2025
Accepted August 25, 2025
Published August 30, 2025

Keywords
Diabetes Prediction;
Feedforward Neural Network;
Adam Optimizer;
SMOTE;
Medical Classification;
Machine Learning;
Imbalanced Data

Author Email
arizhawk21@gmail.com
mazdadi@ulm.ac.id
dwikartini@ulm.ac.id
andifarmadi@ulm.ac.id
f.indriani@ulm.ac.id

I. Introduction

Diabetes mellitus is a chronic metabolic disorder characterized by elevated blood glucose levels due to insulin deficiency or resistance. According to the World Health Organization (WHO), more than 422 million people worldwide live with diabetes, the majority of whom are in low- and middle-income countries. This number is projected to rise significantly in the coming years[1][2]. If left untreated, diabetes can lead to serious complications such as kidney failure, cardiovascular disease, and neuropathyb[3]. However, early diagnosis still relies on conventional lab-based tests, such as blood glucose and oral glucose tolerance tests, which can be time-consuming and resource-intensive[4][5].

To address these challenges, artificial intelligence particularly machine learning (ML) offers a promising alternative for fast and accurate disease prediction. ML enables systems to learn from past experiences and recognize patterns in data without explicit programming[6][7]. It has proven effective in analyzing medical datasets, such as the Pima Indian Diabetes dataset, and extracting predictive insights. Studies have shown that ML techniques can significantly enhance the diagnostic process in healthcare, allowing for early detection and better risk management[8]. For example, research by J. P. Kandhasamy & S. Balamurali[9] demonstrated how various ML algorithms can predict diabetes. Additionally, studies by H. Setiawan, A. Firnanda & U. Khair[10] confirmed that FNN could achieve accurate predictions for diabetes, highlighting the potential of neural networks in medical classifications. One widely used approach in medical classification is the Feedforward Neural Network (FNN), a type of neural network where data flows in one direction from input to output, making it suitable for modeling nonlinear relationships[11].

However, its performance depends heavily on optimal parameter settings, such as weights and biases, which

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127
Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

must be adjusted to achieve reliable predictions[12]. For example, optimizing these parameters using techniques like Adaptive Moment Estimation (ADAM) has been shown to improve the accuracy of models significantly. Kingma & Ba[13] highlighted the effectiveness of ADAM in improving convergence and training stability, making it a go-to optimization technique in neural network models. Furthermore, research by Alex et al.[14] and Shirley & Priya[15] have demonstrated that techniques like SMOTE can address class imbalance, ensuring that the model is trained on a more balanced dataset, thus enhancing the model's ability to generalize and reduce biases.

Although numerous studies have demonstrated the effectiveness of ML models in predicting diabetes, most research has either not optimized neural networks with advanced techniques such as ADAM or overlooked the issue of class imbalance using methods like SMOTE[16]. Furthermore, few studies have combined both ADAM optimization and SMOTE to improve the performance of FNN in predicting diabetes, particularly in real-world medical datasets[17]. This gap presents an opportunity for further research to explore the combined effects of these techniques on improving diabetes prediction models.

This research addresses the limitations identified in previous studies by combining the FNN model with ADAM optimization and the SMOTE technique. By optimizing the FNN model using ADAM, we can fine-tune its parameters, leading to better generalization and more accurate predictions. SMOTE will be applied to handle class imbalance, ensuring that the model learns effectively from the minority diabetic class. Combining these techniques is expected to significantly enhance the accuracy of diabetes prediction models, especially in datasets with class imbalances.

Therefore, the primary objective of this research is to evaluate the performance of an FNN model optimized with ADAM and enhanced with SMOTE to predict the early stages of diabetes. This study will compare the performance of the optimized model against an FNN model without optimization to demonstrate the improvements in prediction accuracy.

This research makes several key contributions to the field of diabetes prediction. Firstly, it improves diabetes prediction accuracy by combining ADAM optimization with Feedforward Neural Networks (FNN) and addressing class imbalance by applying the Synthetic Minority Oversampling Technique (SMOTE). This approach aims to enhance the model's ability to make more accurate predictions by effectively training on a balanced dataset. Secondly, the study explores the efficiency of using ADAM and SMOTE in optimizing and balancing FNN models for medical applications, showcasing their potential to improve performance in complex healthcare datasets. The practical implications of this research are significant, as the proposed method can assist in early diabetes detection, enabling clinicians to make more informed and timely decisions that could improve patient outcomes. Lastly, this research offers a comprehensive comparison between a baseline FNN and an optimized FNN, providing valuable insights into the impact of optimization techniques on model performance and highlighting the importance of incorporating advanced optimization methods for medical predictive tasks.

This paper is structured as follows: Section 1 discusses the related work and previous studies in diabetes prediction using Feedforward Neural Networks (FNN). Section 2 presents the methodology and SMOTE technique, detailing the FNN model, ADAM optimization, and the dataset used. Section 3 discusses the experimental results and performance evaluation. Finally, Section 4 concludes the paper and suggests directions for future research.

II. Materials And Method

A. Data Collection

This study utilized the publicly available diabetes dataset provided by the *National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)* via Kaggle https://www.kaggle.com/datasets/mathchi/diabetes-data-set. The dataset is widely used in medical classification research and comprises 768 samples with nine attributes including eight input features and one binary target label (Outcome)[18]. The characteristics of the dataset used in this study, including the feature names, data types, units, and descriptions, are summarized in Table 1.

Table 1. Description of Dataset Features.

No.	Feature	Data Type	Unit	Description
1	Pregnincies	Integer	-	Number of pregnancies experienced by the female patient
2	Glucose	Integer	mg/dL	Plasma glucose concentration 2 hours after glucose intake
3	Blood Pressure	Integer	mmHg	Diastolic blood pressure
4	Skin Thickness	Integer	Mm	Triceps skinfold thickness
5	Insulin	Integer	μU/mL	Serum insulin level 2 hours after glucose intake
6	BMI	Float	kg/m^2	Body Mass Index

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127
Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

7	Diabetes Function	Float	-	Genetic diabetes likelihood score based on family history
8	Age	Integer	Years	Patient age at the time of data collection
9	Outcome	Category	0/1	Target class 0 = non-diabetic, 1 = diabetic

B. Data Preprocessing

Data preprocessing was carried out in several steps to ensure high-quality inputs for the neural network model:

1. Invalid Value Handling

Several features (e.g., Glucose, Blood Pressure, Skin Thickness, and Insulin) contained zero clinically unrealistic values. These were treated as *missing values* and imputed using the mean of the respective feature with the formula Eq. (1)[19]:

$$x_i = \begin{Bmatrix} x_i, if x_i \neq 0 \\ \bar{x}, if x_i = 0 \end{Bmatrix} \quad (1)$$

The formula defines x_i as a piecewise function where x_i remains unchanged if non-zero, and is replaced by $\bar{x}$(such as mean or constant) if $x_i = 0$, to avoid distortion in data analysis. This method preserves the overall distribution and avoids bias[19].

2. Feature Normalization

As the features have different scales, Min-Max normalization was applied Eq. (2)[20]:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

where x' is the normalized value, x is the original value, and x_{min} and x_{max} are the minimum and maximum values in the dataset. This transformation standardizes input values into the range [0, 1] and helps improve model convergence and stability[21][22].

C. Data Splitting

The input data underwent a normalization process that standardized its values to the range [0, 1], a step essential for improving the convergence speed and stability of the model during training[23]. After this preprocessing stage, the dataset was partitioned into two subsets through stratified sampling, ensuring that the proportion of diabetic and non-diabetic samples remained consistent across both sets. Specifically, 80% of the data was allocated for training purposes, while the remaining 20% was reserved for testing[24]. This stratification technique preserves the original class distribution within each subset, crucial for maintaining the model's performance across imbalanced classes[25].

D. Synthetic Minority Oversampling Technique (SMOTE)

An initial class imbalance was observed, where only about 35% of the data belonged to the diabetic class. To address this, the Synthetic Minority Oversampling Technique (SMOTE) was applied. SMOTE generates synthetic data points for the minority class by interpolating between existing minority samples and their nearest neighbors in the feature space.

The SMOTE formula is defined as Eq. (3)[26]:

$$x_{new} = x_i + \delta \cdot (x_{zi} - x_i) \quad (3)$$

The term x_{new} represents the updated value of the variable, while x_i is the original value of the variable in question, and $x_{zi}$is the target or reference value that we aim to approach. The factor δ determines the degree of change that occurs between x_i and x_{zi}. When $\delta = 0$, the new value x_{new} will remain the same as the original value x_i, indicating no change. Conversely, when $\delta = 1$, the new value will equal the target value x_{zi}, meaning the change is complete. SMOTE effectively enhances minority class representation while reducing the risk of overfitting by avoiding duplicate samples[27][26].

E. Feedforward Neural Network (FNN)

This study utilized a Feedforward Neural Network (FNN) model architecture. FNN was chosen due to its ability to model complex nonlinear relationships between features, which is crucial in medical classification. The architecture is structured as follows in Table 2:

Table 2. Architecture of the Proposed Neural Network Model.

Layer	Number of Neurons	Activation Function
Input Layer	8	-
Hidden Layer 1	16	ReLU
Output Layer	1	Sigmoid

Each neuron in the network computes a weighted sum of its inputs, followed by the application of an activation function Eq. (4)[28]:

$$z = \sum_{i=1}^{n} w_i x_i + b \quad (4)$$

In this formula, z is the output of the function. w_i represents the weights or parameters that control the contribution of each input x_i to the output. The variable x_i is the input value for each feature or variable in the dataset, and b is the bias term, allowing for shifting the linear function's result. In the hidden layers, the Rectified Linear Unit (ReLU) function is applied Eq. (5)[29]:

$$ReLU(z) = max(0, z) \quad (5)$$

In this formula, $ReLU(z)$ represents the output of the activation function, while z is the input to the function, typically the result of a linear transformation from the

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

previous layer's outputs in a neural network. The expression $\max(0, z)$ means that the function will return the maximum value between 0 and z. In other words, if z is negative, the output will be 0; if z is positive, the output wll be z itself. ReLU is chosen for its computational efficiency and ability to mitigate the vanishing gradient problem in deep networks[29]. For the output layer, a sigmoid activation function is applied Eq. (6)[28]:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{6}$$

In this equation, $\sigma(z)$ is the output the sigmoid function, which produces a value between 0 and 1, often used as a probability in various applications. The variable z is the input to the sigmoid function, typically a linear combination of features, such as $z = w_1 x_1 + w_2 x_2 + b$, where w_1,w_2 are the weights and b is the bias. The constant e is Eular's number, which is used to exponentiate the negative value of z. This function squashes the output into a probability range [0,1], ideal for binary classification tasks such as diabetes prediction[28]. The output layer consists of a single neuron that generates a scalar output interpreted as the probability of a patient having diabetes. The output $\hat{y}$ is compared to the true label $y \epsilon \{0,1\}$using the binary cross-entropy loss function Eq. (7)[30]:

$$L(y, \hat{y}) = -\big[y \cdot \log(\hat{y}) + (1 - y) \cdot \log\big(1 - (\hat{y})\big)\big] \tag{7}$$

In this formula, $L(y, \hat{y})$ represents the loss or error value, quantifying the difference between the actual label y and the predicted probability $\hat{y}$ provided by yhe model. y is the true label, which takes values of either 0 or 1, indicating the class of the input data. On the other hand, $\hat{y}$ is the probability predicted by the model for class 1. The logarithmic function $\log(\hat{y})$ calculates the contribution from class 1, and $\log(1 - \hat{y})$ computes the contribution from class 0. The factor $0 - y$ ensures that the calculation considers the contribution from class 0 when the true label is 0. This loss function penalizes incorrect classifications and is standard in probabilistic binary classification[15][30].

F. Adaptive Moment Estimation (ADAM)

The Adam (Adaptive Moment Estimation) algorithm was applied to optimize the learning process. Adam combines the benefits of momentum and RMSProp to adaptively update model weights[13]. The weight update rule is defined as Eq. (8)[31].

$$\theta_{t+1} = \theta_t - \frac{\alpha \cdot \hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{8}$$

The given formula defines the relationship between the current state and the next state of the variable θ in an iterative process. Here, θ_{t+1} represents the next value of θ at time $t + 1$, while θ_t is the current value of θ at time t. The term α is a constant, often referred to as the learning rate, which controls how much the other components influence the update. $\hat{m}_t$ is a variable that typically represents a moving average or momentum term, indicating how much influence past values should have on the current update. The denominator includes $\hat{v}_t$, which is another variable, likely related to variance or another form smoothing. finally ϵ is a small constant, often used to avoid division by zero and stabilize the computation.

By employing Adam, the model benefits from fast convergence, effective learning rate adaptation, and resilience to noisy gradients. These characteristics are crucial in biomedical applications where datasets may not always be clean or balanced, and where the model's predictive performance can directly influence critical clinical decisions[31].

G. Confusion Matrix

The performance of the proposed Feedforward Neural Network (FNN) was evaluated using a confusion matrix-based approach. This matrix compares predicted and actual labels to provide a structured assessment of classification accuracy. Since the dataset is imbalanced, relying solely on accuracy may not reflect the model's effectiveness. Therefore, additional metrics precision, recall, and F1-score were calculated to offer a more comprehensive evaluation[24]. These metrics were computed using the following formulas Eq. (9), Eq. (10), Eq. (11), and Eq (12) [32][33]:

- Accuracy :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

- Precision :

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

- Recall :

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

- F1-Score :

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{12}$$

The symbols used in the classification performance metrics are defined as follows: TP (True Positive) refers to the number of correctly predicted positive instances, TN (True Negative) denotes the number of correctly predicted negative instances, FP (False Positive) represents the number of negative instances incorrectly predicted as positive, and FN (False Negative) indicates the number of positive instances incorrectly predicted as negative. These formulas are particularly critical in medical classification tasks where the cost of misclassification can have significant clinical implications. These metrics together allow for a balanced evaluation of the model's predictive ability, especially in a medical context where misclassifications can carry serious consequences. A confusion matrix heatmap was also generated to visualize prediction outcomes and identify potential class bias[33][34].

III. Results

A. SMOTE Process

To address the class imbalance in the PIMA Indian Diabetes dataset, the Synthetic Minority Over-sampling

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Technique (SMOTE) was employed during the data preprocessing stage. Initially, the dataset exhibited a significant disparity between the number of non-diabetic (Class 0) and diabetic (Class 1) samples. Such an imbalance tends to bias classification models toward the majority class, leading to poor generalization and reduced sensitivity in identifying minority class instances critical in medical diagnoses such as diabetes.

SMOTE was applied to the training data to synthetically generate new instances of the minority class by interpolating between neighboring samples. This technique avoids simple duplication and introduces more diverse minority class examples, helping the model learn a better decision boundary.

The effectiveness of SMOTE is demonstrated in Fig. 1, which visualizes the class distribution before and after applying the technique. Before SMOTE, Class 1 (diabetic) was significantly underrepresented, while Class 0 (non-diabetic) dominated the dataset. After applying SMOTE, the class distribution became balanced, providing equal representation of both classes. This rebalancing is crucial to ensure that the classifier does not become biased toward the majority class and is capable of detecting diabetic cases more reliably.

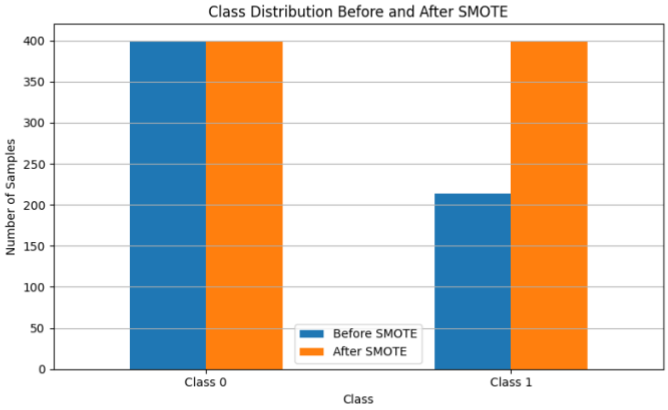

Fig. 1. Class distribution before and after applying SMOTE.

B. Model Performances

To evaluate the impact of optimization techniques on the performance of Feedforward Neural Networks (FNN) in classifying diabetes cases, two model configurations were implemented: an unoptimized FNN trained manually using basic gradient descent, and an optimized FNN trained using the Adam optimizer. Both models underwent identical preprocessing procedures, including zero-value imputation using the mean, feature scaling with StandardScaler, and class balancing through SMOTE (Synthetic Minority Oversampling Technique). Model evaluation was conducted on the same test dataset using standard classification metrics: accuracy, precision, recall, and F1-score.

The unoptimized FNN was implemented manually with a single hidden layer consisting of 16 neurons and ReLU activation, followed by a sigmoid-activated output layer. The model was trained using full-batch gradient descent with a learning rate of 0.01 for 900 epochs. The evaluation yielded an accuracy of 0.7013, a precision of 0.5606, a recall of 0.6852, and an F1-score of 0.6167. These results indicate that while the model showed reasonably good sensitivity in detecting diabetes cases (as seen in the high recall), its relatively low precision suggests many false positives, reducing its reliability for clinical applications.

In comparison, the second FNN model employed the Adam optimizer within the Keras framework, but maintained the same architecture and training settings as the manual model. The training was repeated across ten independent runs to account for stochastic effects arising from weight initialization and adaptive optimization. The average performance across these trials showed consistent improvements, with an accuracy of 0.7331, a precision of 0.6097, a recall of 0.6667, and an F1-score of 0.6364. Compared to the unoptimized FNN, the Adam-optimized model demonstrated improvements in all metrics except for a slight decrease in recall. However, the increase in precision led to a better overall classification balance, as reflected in the higher F1-score. These findings suggest that the Adam optimizer enhances the reliability and robustness of FNN models in the context of medical diagnosis, particularly for predicting diabetes based on clinical data.

Fig. 2, illustrates the comparison of accuracy, precision, recall, and F1-score between the unoptimized FNN model and the FNN model trained using the Adam optimizer The optimized model consistently outperformed the unoptimized version across most metrics, particularly in precision and F1-score, confirming the positive impact of adaptive optimization in neural network training for medical classification tasks.

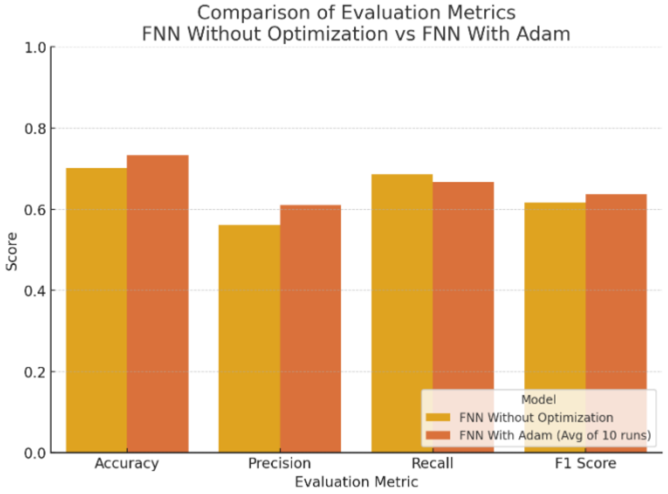

Fig. 2. Comparison of Evaluation Metrics Between FNN Without Optimization and FNN With Adam Optimizer.

Table 3 summarizes the classification performance of two FNN models: one without optimization and one using the Adam optimizer. The unoptimized model achieved an accuracy of 0.7013 and an F1-score of 0.6167, while the Adam-optimized model, averaged over ten runs,

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

improved these metrics to 0.7331 and 0.6364, respectively. Although the recall slightly decreased, the increase in precision led to a more balanced and overall better-performing model. These results highlight the Adam optimizer's positive impact in enhancing neural network performance for diabetes prediction.

Table 3. Comparison of Classification Performance Between Unoptimized and Adam-Optimized FNN Models.

Model	Accuracy	Precision	Recall	F1 Score
FNN (Manual, No Optimizer)	0.7013	0.5606	0.6852	0.6167
FNN + Adam (Avg of 10 runs)	0.7331	0.6097	0.6667	0.6364

IV. Discussion

The present study implemented a Feedforward Neural Network (FNN) optimized with the Adam optimizer to address the challenge of diabetes prediction using the PIMA Indians Diabetes Dataset. The modeling process involved several essential preprocessing stages, including missing value imputation using the mean strategy, feature standardization, and class rebalancing through the Synthetic Minority Over-sampling Technique (SMOTE). These preprocessing steps were critical to ensuring fair training and unbiased evaluation of the neural network, particularly given the inherent class imbalance in the original dataset.

In comparison to related works, traditional classification methods such as decision tree-based algorithms, probabilistic approaches, and kernel-based techniques have been widely applied to the same dataset. While these methods offer advantages in simplicity of implementation and interpretability, they often encounter limitations in capturing the complex non-linear patterns embedded within biomedical data. By contrast, neural network architectures such as FNNs possess a greater capacity to approximate complex mappings due to their multi-layered structure and high representational power. When combined with adaptive optimization strategies like Adam, the learning process becomes more effective, enabling faster convergence and reducing the risk of becoming trapped in local minima. The results obtained in this study reflect these strengths, as the proposed FNN–Adam configuration delivered superior predictive performance compared to previously reported methods. To better understand the proposed method, this study compared it with several machine learning algorithms reported in previous works using the Pima Indians Diabetes Dataset. The summary of the comparison is shown in Table 4.

For instance, Saravananathan and Velmurugan[35] compared J48, CART, SVM, and kNN using 10-fold cross-validation, with J48 achieving the highest accuracy at 67.16%, followed by SVM at 65.05%, CART at 62.29%, and kNN at 53.39%. Rajni and Amandeep[36] developed RB-Bayes, a modification of Naïve Bayes that addresses the *zero probability* issue by replacing missing values with the mean. This method achieved the highest accuracy in their study at 72.90%, outperforming SVM (70.90%), Decision Tree (68.18%), and standard Naïve Bayes (67.71%).

Table 4. Comparison of Classification Accuracy with Previous Studies.

Methods	Results
J48[35]	67.15%
SVM[35]	65.05%
CART[35]	62.29%
kNN[35]	53.39%
RB-Bayes[36]	72.90%
KNN (SMOTE)[37]	72.10%
BNB (SMOTE)[37]	69.20%
NB[38]	72.60%
KNN[38]	66.10%
DT[38]	71.80%
RF[38]	64.90%
Kstar[39]	68.23%
oneR[39]	70.83%
SMO[39]	72.14%
SVM[40]	69.00%
NB[40]	67.50%
KNN[40]	70.50%
FNN + Adam (Proposed Method)	73.31%

Meanwhile, O. Iparraguirre-Villanueva et al.[37] examined the effect of SMOTE on classification performance and reported that KNN (SMOTE) achieved 72.10% accuracy and BNB (SMOTE) achieved 69.20%, confirming that class balancing can enhance accuracy in medical datasets. Additional findings from Kangra and Singh[38] showed that Naïve Bayes (72.60%) outperformed Decision Tree (71.80%), Random Forest (64.90%), and KNN in their evaluation. S. Larabi-Marie-Sainte et al.[39] assessed Kstar, SMO, and oneR, with 71.12%, 72.14%, and 68.23%, respectively. While SMO achieved the highest performance among the three, its accuracy was still lower than RB-Bayes[36] and the proposed method.

The proposed FNN + Adam method achieved the highest accuracy at 73.31%, surpassing all compared methods. This demonstrates that combining an appropriate neural network architecture with Adam optimization can effectively capture complex non-linear relationships between features while avoiding local minima. These findings suggest that deep learning approaches with adaptive optimization offer greater

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127
Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

potential than classical methods such as RB-Bayes, Naïve Bayes, or SMO, particularly in the context of disease prediction in the healthcare domain.

Despite promising results, this study has several limitations. First, the evaluation used only a single dataset (Pima Indians Diabetes Dataset) with a limited sample size (768 entries) and class imbalance. Second, the study did not assess the model's generalizability to other medical datasets with different characteristics, such as varying feature counts, missing value rates, or more extreme class imbalance. Third, no direct comparison was made with ensemble deep learning or transfer learning approaches, which might yield even better performance.

The implications of this study are substantial, particularly for the development of clinical decision support systems. The results indicate that applying FNN with adaptive optimization, such as Adam, can significantly improve diabetes prediction accuracy compared to conventional machine learning methods. This could accelerate early screening processes, reduce false negatives, and enable earlier medical interventions. Moreover, this approach can be adapted to predict other diseases using tabular clinical datasets, with minimal architecture and model parameters modifications.

V. Conclusion

This study aimed to evaluate the effectiveness of Feedforward Neural Networks (FNN) in predicting diabetes by comparing a manually trained model and one optimized using the Adam algorithm, both enhanced with SMOTE to address class imbalance. Experimental results showed that the unoptimized FNN achieved an accuracy of 70.13%, precision of 56.06%, recall of 68.52%, and F1-score of 61.67%. In contrast, the Adam-optimized model yielded improved performance with an average accuracy of 73.31%, precision of 60.97%, recall of 66.67%, and F1-score of 63.64% over ten runs, indicating a more balanced and reliable classification. For future work, deeper network architectures, integration of explainable AI techniques, and application to real-time or longitudinal diabetes data are recommended to enhance prediction accuracy and clinical relevance further.

References

[1] S. Simanto, K. Mridha, R. Saha, M. Limbu, A. Ghosh, and R. N. Shaw, "Diabetes Prediction Using Machine Learning Techniques," *Lect. Notes Electr. Eng.*, vol. 914, no. 09, pp. 317–333, 2022, doi: 10.1007/978-981-19-2980-9_26.

[2] W. Wayunah, W. Nur Aeni, M. Kuncoro, and M. Saefulloh, "Decreased Foot Sensitivity as a Predictor of Diabetic Foot Ulcers," *J. VNUS (Vocational Nurs. Sci.*, vol. 6, no. 1, pp. 1–8, 2024, doi: 10.52221/jvnus.v6i1.550.

[3] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Deep Learning for Diabetes: A Systematic Review," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 7, pp. 2744–2757, 2021, doi: 10.1109/JBHI.2020.3040225.

[4] D. Y. Utami, E. Nurlelah, and F. N. Hasan, "Comparison of Neural Network Algorithms, Naive Bayes and Logistic Regression to predict diabetes," *J. Informatics Telecommun. Eng.*, vol. 5, no. 1, pp. 53–64, 2021, doi: 10.31289/jite.v5i1.5201.

[5] M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers," *IEEE Access*, vol. 8, pp. 76516–76531, 2020, doi: 10.1109/ACCESS.2020.2989857.

[6] T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, "Deep neural networks to predict diabetic retinopathy," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 5, pp. 5407–5420, 2023, doi: 10.1007/s12652-020-01963-7.

[7] A. Mujumdar and V. Vaidehi, "Diabetes Prediction using Machine Learning Algorithms," *Procedia Comput. Sci.*, vol. 165, pp. 292–299, 2019, doi: 10.1016/j.procs.2020.01.047.

[8] S. Sarkar and C. James C. L., "Machine Learning in Healthcare Communication," *Encyclopedia*, pp. 220–239, 2021.

[9] J. P. Kandhasamy and S. Balamurali, "Performance analysis of classifier models to predict diabetes mellitus," *Procedia Comput. Sci.*, vol. 47, no. C, pp. 45–51, 2015, doi: 10.1016/j.procs.2015.03.182.

[10] H. Setiawan, A. Firnanda, and U. Khair, "Enhancing the Accuracy of Diabetes Prediction Using Feedforward Neural Networks: Strategies for Improved Recall and Generalization," *Brill. Res. Artif. Intell.*, vol. 4, no. 1, pp. 201–207, 2024, doi: 10.47709/brilliance.v4i1.3888.

[11] D. Wang and T. Fang, "Forecasting Crude Oil Prices with a WT-FNN Model," *Energies*, vol. 15, no. 6, 2022, doi: 10.3390/en15061955.

[12] R. Chatterjee *et al.*, "FNN for Diabetic Prediction Using Oppositional Whale Optimization Algorithm," *IEEE Access*, vol. 12, no. January, pp. 20396–20408, 2024, doi: 10.1109/ACCESS.2024.3357993.

[13] D. P. Kingma and J. L. Ba, "ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION," pp. 1–15, 2015.

[14] S. A. Alex, N. Z. Jhanjhi, M. Humayun, A. O. Ibrahim, and A. W. Abulfaraj, "Deep LSTM Model for Diabetes Prediction with Class Balancing by SMOTE," *Electron.*, vol. 11, no. 17, 2022, doi: 10.3390/electronics11172737.

[15] J. J. Shirley and M. Priya, "An Adaptive Intrusion Detection System for Evolving IoT Threats: An Autoencoder-FNN Fusion," *IEEE Access*, vol. 13, no. January, pp. 4201–4217, 2025, doi: 10.1109/ACCESS.2024.3525074.

[16] M. Pradhan and R. Kumar Sahu, "Predict the onset of diabetes disease using Artificial Neural Network (ANN)," *Int. J. Comput. Sci. Emerg.*

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Technol. (E, vol. 2, no. 2, pp. 2044–6004, 2011.

[17] D. Dablain, B. Krawczyk, and N. V. Chawla, "DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 34, no. 9, pp. 6390–6404, 2023, doi: 10.1109/TNNLS.2021.3136503.

[18] J. J. Khanam and S. Y. Foo, "A comparison of machine learning algorithms for diabetes prediction," *ICT Express*, vol. 7, no. 4, pp. 432–439, 2021, doi: 10.1016/j.icte.2021.02.004.

[19] X. Feng, Y. Cai, and R. Xin, "Optimizing diabetes classification with a machine learning-based framework," *BMC Bioinformatics*, vol. 24, no. 1, pp. 1–20, 2023, doi: 10.1186/s12859-023-05467-x.

[20] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Appl. Soft Comput.*, vol. 97, no. xxxx, p. 105524, 2020, doi: 10.1016/j.asoc.2019.105524.

[21] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K. R. Müller, "Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications," *Proc. IEEE*, vol. 109, no. 3, pp. 247–278, 2021, doi: 10.1109/JPROC.2021.3060483.

[22] Q. U. A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. S. Alouini, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020, doi: 10.1109/TWC.2020.2986438.

[23] A. Tolic, B. Mileva Boshkoska, and S. Skansi, "Chrono Initialized LSTM Networks With Layer Normalization," *IEEE Access*, vol. 12, no. August, pp. 115219–115236, 2024, doi: 10.1109/ACCESS.2024.3445329.

[24] M. K. Uçar, M. Nour, H. Sindi, and K. Polat, "The Effect of Training and Testing Process on Machine Learning in Biomedical Datasets," *Math. Probl. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/2836236.

[25] R. Rahim, "Application of the forward selection strategy to the C4.5 Algorithm to improve the classification's accuracy of a breast cancer data set," *J. Popul. Ther. Clin. Pharmacol.*, vol. 30, no. 1, pp. 14–23, Jan. 2023, doi: 10.47750/jptcp.2023.1002.

[26] M. K. Rezki, M. I. Mazdadi, F. Indriani, Muliadi, T. H. Saragih, and V. A. Athavale, "Application of Smote to Address Class Imbalance in Diabetes Disease Categorization Utilizing C5.0, Random Forest, and Support Vector Machine," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 6, no. 4, pp. 343–354, 2024, doi: 10.35882/jeeemi.v6i4.434.

[27] A. Ishaq *et al.*, "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.

[28] A. A. Zulfikar and W. A. Kusuma, "Modeling and predicting protein-protein interactions of type 2 diabetes mellitus using feedforward neural networks," *2019 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2019*, pp. 163–168, 2019, doi: 10.1109/ICACSIS47736.2019.8979989.

[29] H. Zhou, R. Myrzashova, and R. Zheng, "Diabetes prediction model based on an enhanced deep neural network," *Eurasip J. Wirel. Commun. Netw.*, vol. 2020, no. 1, 2020, doi: 10.1186/s13638-020-01765-7.

[30] D. Darbar and I. Bhattacharya, "Application of Machine Learning in Battery: State of Charge Estimation Using Feed Forward Neural Network for Sodium-Ion Battery," *Electrochem*, vol. 3, no. 1, pp. 42–57, 2022, doi: 10.3390/electrochem3010003.

[31] S. Bock and M. Weis, "A Proof of Local Convergence for the Adam Optimizer," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2019-July, no. July, pp. 1–8, 2019, doi: 10.1109/IJCNN.2019.8852239.

[32] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.

[33] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.

[34] O. R. Olaniran, A. R. R. Alzahrani, and M. R. Alzahrani, "Eigenvalue Distributions in Random Confusion Matrices: Applications to Machine Learning Evaluation," *Mathematics*, vol. 12, no. 10, pp. 1–14, 2024, doi: 10.3390/math12101425.

[35] K. Saravananathan and T. Velmurugan, "Analyzing Diabetic Data using Classification Algorithms in Data Mining," *Indian J. Sci. Technol.*, vol. 9, no. 43, 2016, doi: 10.17485/ijst/2016/v9i43/93874.

[36] Rajni and Amandeep, "RB-bayes algorithm for the prediction of diabetic in 'PIMA Indian dataset,'" *Int. J. Electr. Comput. Eng.*, vol. 9, no. 6, pp. 4866–4872, 2019, doi: 10.11591/ijece.v9i6.pp4866-4872.

[37] O. Iparraguirre-Villanueva, K. Espinola-Linares, R. O. Flores Castañeda, and M. Cabanillas-Carbonell, "Application of Machine Learning Models for Early Detection and Accurate Classification of Type 2 Diabetes," *Diagnostics*, vol. 13, no. 14, 2023, doi: 10.3390/diagnostics13142383.

[38] K. Kangra and J. Singh, "Comparative analysis of predictive machine learning algorithms for diabetes mellitus," *Bull. Electr. Eng. Informatics*,

---

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.
Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127
Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

vol. 12, no. 3, pp. 1728–1737, 2023, doi: 10.11591/eei.v12i3.4412.

[39] S. Larabi-Marie-Sainte, L. Aburahmah, R. Almohaini, and T. Saba, “Current techniques for diabetes prediction: Review and case study,” *Appl. Sci.*, vol. 9, no. 21, 2019, doi: 10.3390/app9214604.

[40] M. F. Faruque, Asaduzzaman, and I. H. Sarker, “Performance Analysis of Machine Learning Techniques to Predict Diabetes Mellitus,” *2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE 2019*, 2019, doi: 10.1109/ECACE.2019.8679365.

AUTHOR BIOGRAPHY

Arizha Wijaya Kusuma, a Department of Computer Science student at Lambung Mangkurat University, has conducted an innovative research project entitled *“Improving Diabetes Prediction Using Feedforward Neural Network with Adam Optimization and SMOTE Technique.”* This research reflects his strong interest in artificial intelligence and medical informatics, particularly in applying machine learning for health diagnostics. In his study, Arizha Wijaya Kusuma developed a classification model based on the Feedforward Neural Network (FNN). She enhanced it using the Adaptive Moment Estimation (Adam) optimizer alongside the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance in diabetes prediction. The integration of these techniques resulted in improved classification accuracy and robustness, highlighting his capability in data-driven modeling and AI-based healthcare solutions.

Muhammad Itqan Mazdadi, a lecturer in the Department of Computer Science, Lambung Mangkurat University. His research interest is centered on Data Science and Computer Networking. Before becoming a lecturer, he completed his undergraduate program in the Computer Science Department at Lambung Mangkurat University In 2013. He then completed his master’s degree from the Department of Informatics at Islamic University of Indonesia, Yogyakarta. Currently, he serves as the Secretary of the Computer Science Department at Lambung Mangkurat University. Email: mazdadi@ulm.ac.id. Orcid ID: 0000-0002-8710-4616.

Dwi Kartini earned her Bachelor's and Master's degrees in Computer Science from the Faculty of Computer Science at Putra Indonesia "YPTK" in Padang, Indonesia. She is also a lecturer in the Department of Computer Science, where she teaches various subjects including linear algebra, discrete mathematics, research methods, and others. Her research interests focus on the applications of Artificial Intelligence and Data Mining. Currently, she serves as an assistant professor in the Department of Computer Science, Faculty of Mathematics and Natural Sciences at Lambung Mangkurat University in Banjarbaru, Indonesia. She is the head of the Computer Science study program. She can be contacted at dwikartini@ulm.ac.id. Orcid ID: 0000-0002-7382-5084

Andi Farmadi, a senior lecturer in the Computer Science program at Lambung Mangkurat University. He has been teaching since 2008 and has been the Head of the Data Science Lab since 2018. He completed his undergraduate studies at Hasanuddin University and his graduate studies at Bandung Institute of Technology. His research area, up to the present, focuses on Data Science. One of his research projects, along with other researchers, published in the International Conference of Computer and Informatics Engineering (IC2IE), is titled "Hyperparameter tuning using GridsearchCV on the comparison of the activation function of the ELM method to the classification of pneumonia in toddlers," and this research was published in 2021. Email: andifarmadi@ulm.ac.id. Orcid ID: 0009-0009-0926-8082.

Fatma Indriani is a lecturer in the Department of Computer Science at Lambung Mangkurat University, with a strong research interest in Data Science. Before pursuing an academic career, she completed her undergraduate studies in the Informatics Department at the Bandung Institute of Technology. In 2008, she began her journey as a lecturer at Lambung Mangkurat University, contributing to Computer Science through teaching and research.To further expand her expertise, she pursued a master’s degree at Monash University, Australia, which she completed in 2012. Her academic journey continued with a doctorate in Bioinformatics from Kanazawa University, Japan, which she completed in 2022. With a focus on both Data Science and Bioinformatics, she actively engages in research, exploring innovative ways to leverage data-driven technologies for scientific advancement. Her dedication to academia and research allows her to contribute significantly to the development of knowledge in her field while mentoring students and collaborating on interdisciplinary projects. Email: f.indriani@ulm.ac.id Orcid ID: 0009-0006-7180-6708.

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Dr. P. Chandrasekaran is a biochemistry researcher and ICMR Project Assistant at KS Rangasamy College of Arts and Science, Tamil Nadu, India. His work spans therapeutic phytomedicine, metabolic disorders, and molecular pharmacology, with over 20 publications in peer-reviewed journals and book chapters. He has contributed significantly to understanding obesity-related pathophysiology and phytotherapeutic interventions, including bioactive compounds like biochanin A and hydroxycitric acid. Dr. Chandrasekaran is also an active peer reviewer for leading journals in biochemistry and pharmacology. His interdisciplinary approach bridges natural product research and systems biology, advancing innovative strategies for chronic disease management.

Corresponding author: Muhammad Itqan Mazdadi, mazdadi@ulm.ac.id, Department of Computer Science, Lambung Mangkurat University, Jl. A. Yani Km 36, Banjarbaru 70714, Indonesia.

Digital Object Identifier (DOI): https://doi.org/10.35882/ijeeemi.v7i3.127

Copyright © 2025 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).